Intelligent Tutoring Systems: Then and Now

J. Dexter Fletcher
Institute for Defense Analyses
Alexandria, VA

Intelligent Tutoring Systems: Then and Now

J. Dexter Fletcher
Institute for Defense Analyses
1801 North Beauregard Street
Alexandria, VA 22311

This talk was presented on March 10, 1999 by Dr. Dexter Fletcher who is a research staff member at the Institute for Defense Analyses (IDA) in Alexandria, Virginia. IDA's sole function is to perform studies and analyses of scientific and technical matters for the Office of the Secretary of Defense (OSD). As usual, this presentation represents the views of the presenter and does not represent official policies or positions of either IDA or OSD.

The phrase Intelligent Tutoring Systems (ITS) covers a form of computer-based instruction (CBI) that has also been called intelligent computer-assisted instruction (ICAI). ITSs may be as intelligently or un-intelligently designed as any other form of CBI. 'Intelligent' in this case refers to a particular functionality that is the goal of these systems and is further defined in this discussion.

Although much in this presentation may be relevant to private and public sector education, it is focused on applications of ITS to military training. The presentation is also fairly compressed. More could be said about all the issues it raises.

Intelligent Tutoring Systems: Then and Now

Workshop on Advanced Training Technologies and Learning Environments

9-10 March 1999

J. D. Fletcher
Institute for Defense Analyses

Summary

The object of this presentation is to suggest that:

Substantial improvements in instructional effectiveness may be obtained through increased tailoring of military training to the needs and capabilities of individual learners. Routine provision of a single human instructor for every student is, with a few significant exceptions, prohibitively expensive.

Individualization of military training and its substantial benefits can be significantly increased through the use of ITSs which generate instructional interactions tailored to the needs of individual students. They substitute technology for human labor, thereby making increased individualization economically feasible.

ITSs can thereby make high-quality training available any time, any place and tailored to the needs of any student.

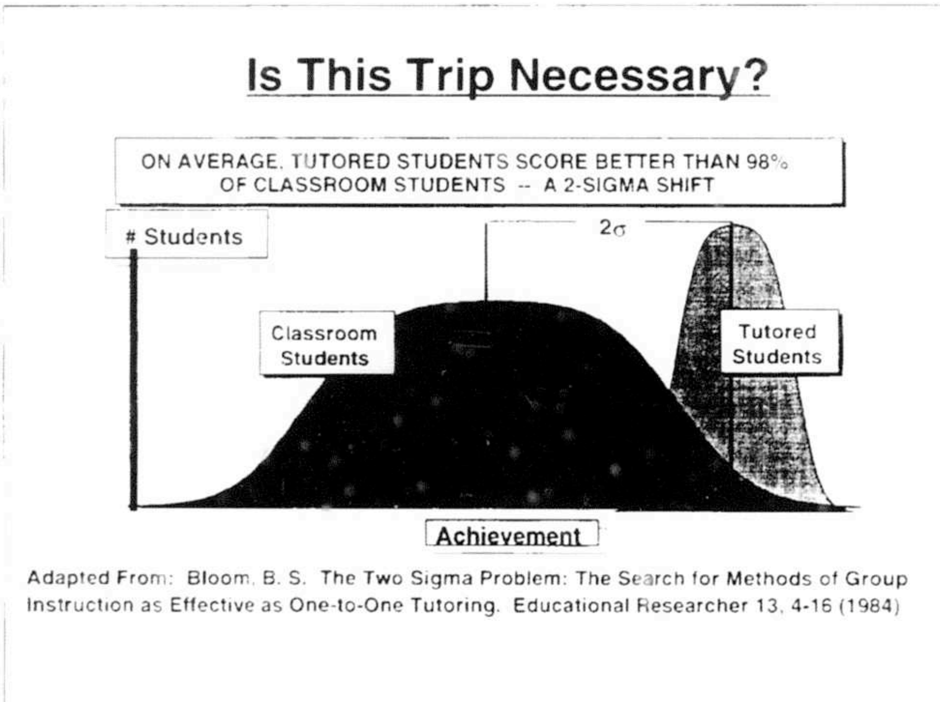
There are no unmanned systems, nor automatic operations. By increasing the accessibility of high-quality, individualized training whenever, wherever, and to whomever it is needed, ITSs can improve readiness and operational effectiveness.

In Short ...

- **Individualization is a good thing, but expensive.**
- **ITSs increase individualization through real-time, on-demand generation of relevant instructional interactions.**
- **ITSs thereby make DoD training more accessible and relevant than is otherwise possible.**
- **ITSs will improve operational effectiveness.**

Individualized Instruction

Why should we care about tailoring training to individual students? Why should we care about individualized instruction? Benjamin Bloom and his students attempted to answer these questions by comparing the achievement of individually tutored students (one (human) teacher per one student) with that of classroom students (one (human) teacher per 28-32 students) (Bloom 1984). They found the difference in achievement to be about two standard deviations, which means, roughly, that tutoring improved the achievement of 50th percentile students to that of 98th percentile students.



Technology-Based Training May Answer Bloom's Challenge

Comments on Bloom's research might include the following:

Intuitively we would expect tutoring to effect some improvement in achievement or learning. What is remarkable about Bloom's finding is the size of the difference. Two standard deviations is a very large difference.

Education is a worthy end in its own right. It is preparation for life and its learning objectives can be more flexible and negotiable than they are in training. Educators are likely to emphasize the maximization of learning assuming fixed costs. Training is a means to an end -- operational effectiveness or readiness in the military and perhaps productivity in the private sector. Its learning objectives are determined by the needs of jobs and careers and not by the variable needs of students. Trainers in both the public and private sectors, therefore, are likely to emphasize the minimization of costs to accomplish fixed levels of learning. The ultimate goal of DoD training is improved readiness and operational effectiveness.

Bloom's research indicates an instructional imperative and an economic impossibility. We need to break out of this dilemma. Technology and specifically ITSs may provide that.

Applications of computer technology are in their infancy. ITSs most probably apply an old metaphor (one-on-one tutoring) to a new technology just as we originally attempted to make carriages run without horses or telegraphs operate without wires. The technology-based instruction analog to the automobile, radio and television has yet to emerge and may not even have been imagined, but it may well be in the direction of intelligent tutoring systems.

Some Comments on Bloom's Theme

- **It is a big difference.**
- **We (DoD) are in this business to increase accessibility and operational effectiveness.**
- **We face a training imperative and an economic impossibility.**
- **What we seek will not be what we get.**

Question: What do ITSs have to offer that we can't get otherwise?

The Challenges of Classroom Instruction: Pace

At this point it seems reasonable to ask what ITSs offer that might solve Bloom's two standard deviation challenge, or as Bloom calls it, the two sigma problem - sigma being a common symbol for standard deviation.

The variability that classroom teachers must accommodate is daunting:

The ratio of time needed by the slowest Kindergarten students to build words from letters to the fastest Kindergarten students was found by Suppes (1964) to be 13:1.

Similar ratios for grade 5 students to master a unit of social studies were found to be 3:1 and 5:1 (Gettinger and White 1980).

In two studies of rates of learning by hearing impaired and native American students the ratios were found to be 4:1 (Suppes, Fletcher and Zanotti 1975 and 1976).

Based on a range of research findings, Carroll (1970) estimated the overall ratio for elementary school students to be 5:1.

Even among highly selected students at a major research university the ratio needed by undergraduates has been found to be 7:1 (Corbett 1998, personal communication).

NB: Individual Differences in Pace

- **Ratio of time needed to build words from letters in grade K -- 13:1**
- **Ratios of time needed to learn in grade 5 -- 3:1 and 5:1**
- **Ratios of time needed by hearing impaired and native American students to reach mathematics objectives -- 4:1**
- **Overall ratio of time needed to learn, K-8 -- 5:1**
- **Ratio of time needed by college undergraduates to learn LISP -- 7:1**

The Challenges of Classroom Instruction: Interactivity

Those who study classroom interactions have found that groups of students ask about three questions an hour and that any single student in a class asks about 0.11 questions an hour. By contrast, students in individual tutorial sessions have been found to ask 20-30 questions an hour and have been required to answer 117-146 questions an hour. Finally, students taking drill and practice computer-based instruction answer 8-12 questions a minute - questions that have been especially selected to meet the needs of the individual student and that are immediately graded.

It is not difficult to see why the intensity of tutorial settings, either based on human tutoring or computer tutoring might produce large differences in achievement, or reach instructional objectives substantially more efficiently.

NB: Classroom and Tutorial Interactivity

- **Average number of questions asked by a class during a classroom hour: 3**
- **Average number of questions asked by any student during a classroom hour -- 0.11**
- **Average number of questions asked by a student during a tutorial hour -- 21.1 (Research methods); 32.2 (Algebra)**
- **Average number of questions students answered during a tutorial hour -- 117.2 (Research Methods); 146.4 (Algebra)**
- **Average number of questions answered per minute in mathematics drill and practice CBI -- 10**

What Individualization Has Technology Provided in the Past?

What about drill and practice? What about this technique of computer assisted instruction that has been used from the late 1950s? What level of tutorial instruction was it able to provide?

Notably, it could (a) accommodate each student's rate of progress, allowing as much or as little time as each individual student needed to reach instructional objectives; (b) adjust the sequence of content to each student's needs; (c) adjust the content itself - different students would receive different content; (d) make the instruction as easy or as difficult as necessary; (e) adjust to the learning style (e.g., verbal versus visual) that was most appropriate for each student. All of these capabilities have been available and used in computer-based instruction since its inception in the 1950s (e.g., Atkinson and Fletcher 1972; Suppes and Morningstar 1972; and Fletcher and Rockway 1986).

<u>What About That Drill and Practice?</u>

It tailored instruction to individuals by adjusting:

- **Rate of progress**
- **Sequence of content**
- **Content itself**
- **Difficulty**
- **Style**

What Are the Unique Contributions of Intelligent Tutoring Systems?

What then is left for ITSs to provide? What can we get from them that is not otherwise available? Three candidate functionalities may deserve mention:

1) The ability to capture the interactions of one-on-one tutoring.

2) The ability to generate instructional material and interactions as needed rather than foresee and pre-store all such materials and interactions needed to meet all possible eventualities.

3) Related to the above, the ability to allow either the computer or the student to ask open-ended questions and initiate instructional dialogue as needed or desired.

The key defining characteristic of ITSs then, is not application of computer techniques from artificial intelligence or knowledge representation, although these may be essential, but rather the functional capability to generate in real-time and on-demand instructional interactions that are tailored to student requests and/or needs. It was this generative capability that motivated DoD to invest in ITSs in the first place (Fletcher 1988). The motivation was to reduce or eliminate the high costs of foreseeing or predicting all possibly needed materials and interactions and then programming them into computer-based instruction.

What, Then, Do We Get From ICAI/ITS?
What's New?

Perhaps three things:

- **Individualized Tutoring**
  (versus Automated Teaching)

- **Generative Capability**
  (versus Pre-stored Material)

- **Mixed-Initiative Dialogue**
  (versus Computer Initiated Dialogue)

The key may be real-time, on-demand generation of instructional interactions tailored to student requests and/or needs.

An Example of An Intelligent Tutoring System At Work

An example of these capabilities at work may be found in a student-computer mixed-initiative dialogue taken from the SOPHIE (Sophisticated Instructional Environment) system (Brown, Burton and DeKleer 1982). SOPHIE used a general-purpose electronic simulator to represent an electronic power supply and create a reactive, tutorial environment for training electronic troubleshooting. SOPHIE employed a semantic network of information about the power supply, one of the most sophisticated natural language understanders ever built into an ITS, and an articulate, expert troubleshooting capability that could not only solve arbitrary faults introduced by students or instructors, but also explain to the student how it solved them.

The reactive environment in SOPHIE responded in a veridical fashion to the actions and decisions of a student engaged in practical problem solving. It employed strategies akin to those used by the student to make inferences about the circuit based on the student's measurements. Because the inferencing strategies were like those of the student, SOPHIE could determine which deductions the student was using, construct a model of his/her state of knowledge, and use this model to generate explanations in terms that the student was prepared to understand. SOPHIE highlighted the need for intelligent training systems not only to solve arbitrary problems, but to do so in a way that can be explained to students and be understood by them.

The first student-SOPHIE interaction is a straight-forward question and answer. SOPHIE has introduced a fault into the simulated power supply and the student is trying to locate and repair it.

ITS Example
Troubleshooting Tutorial Dialogue From SOPHIE

•

•

•

Student: What is the base emitter voltage of the voltage limiting transistor?

SOPHIE: The base emitter voltage of Q5 is -.037 volts.

Student: What about for the current limiting transistor?

SOPHIE: The base emitter voltage of Q8 is .349 volts.

An Example of An Intelligent Tutoring System At Work (Cont'd.)

The second student-SOPHIE interaction is a little less straight-forward because, in addition to simulating the faulted power supply, SOPHIE must be able to parse the dialogue sufficiently to know that the student is asking for the base emitter voltage of transistor Q8. The meaning of the student's question depends heavily on the context established by his prior question.

The third student-SOPHIE interaction continues its dependence on SOPHIE's understanding of context and additionally requires SOPHIE to extract the answer by simulating both the faulted power supply and one that is operating correctly.

The fourth student-SOPHIE interaction indicates a clear step beyond what Brown et al. described as a knowledgeable system to an intelligent system. SOPHIE has progressed from a knowledgeable parsing of its dialogue with the student and simulation of various states of the power supply to a system exercising tutorial intelligence. It shadowed the student's solution path, modeled the student's troubleshooting hypotheses, determined that they are incorrect, elected to capture the dialogue initiative back from the student, and is undertaking a series of tutorial interactions intended to lead the student back to a more correct approach to the problem.

It is difficult to imagine any way to achieve this level of instructional functionality without a generative capability, which is, then, a defining characteristic of ITSs, distinguishing them from other forms of computer-based instruction.

ITS Example (Cont'd.)
Troubleshooting Tutorial Dialogue From SOPHIE

Student: What should it be?

SOPHIE: In a working circuit, the base emitter voltage of Q8 is .593 volts.

Student: Replace Q8.

SOPHIE: I am going to ask you some questions about how Q8 is faulted. Are any junctions shorted?

•
•
•

What's the Underlying Technical Idea?

Achieving these instructional functionalities requires a technical idea. This idea was first articulated by Jaime Carbonell in 1970. It avoids ad-hoc frame-oriented instructional approaches such as those seen in programmed texts and is best described as "intrinsic programming" (Crowder 1959). This frame-oriented approach is still found ubiquitously in computer-based instruction. In its place, Carbonell recommended an information systems orientation approach or, as we might say today, an approach based on knowledge representation.

What Was the Underlying Technical Idea?

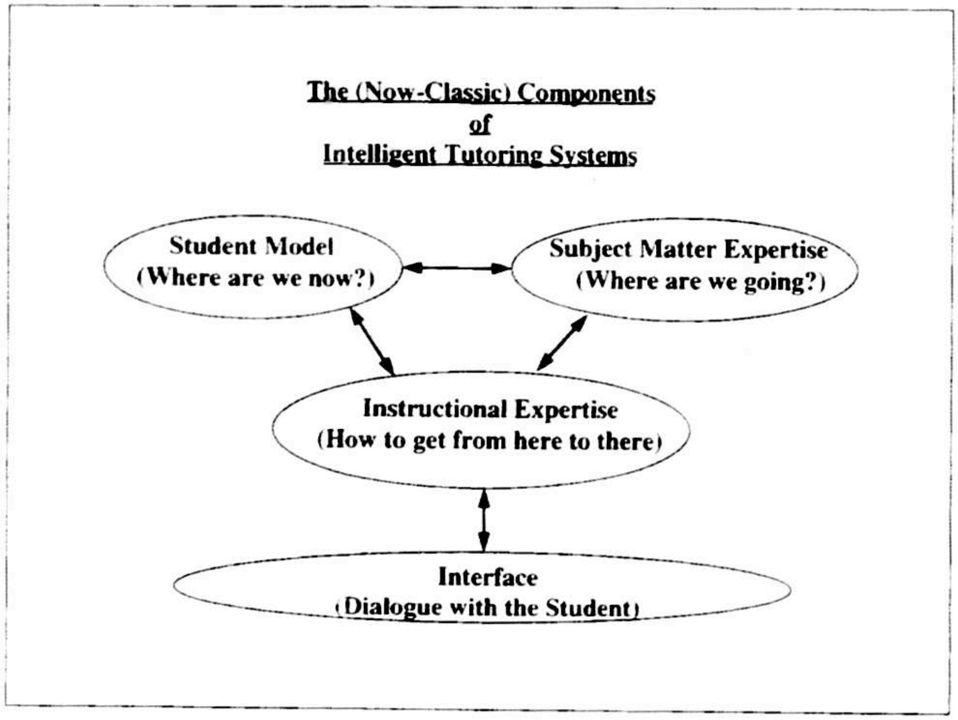
Carbonell, J. R., "AI in CAI: An Artificial Intelligence Approach to Computer-Assisted Instruction," *IEEE Transactions on Man-Machine Systems*, Vol. 11, 1970, pp. 190-202.

The key idea was:

Information Systems Oriented (ISO) Systems
in place of
Ad-hoc Frame Oriented (AFO) Systems

The Structure of Intelligent Tutoring Systems

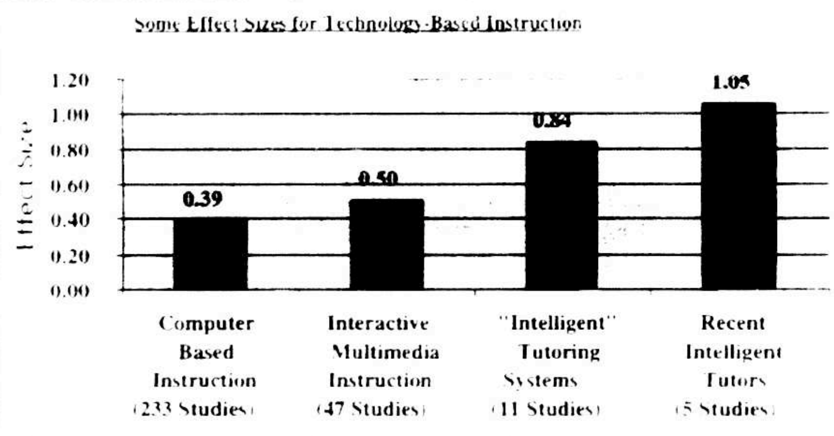
An approach such as that defined by the functionalities possessed by SOPHIE and other ITSs and described by Carbonell involves computer representation or modeling of the student, the subject matter, and expert tutoring (Fletcher 1975). This approach, involving information systems and knowledge representation, is now commonly understood and found to a significant degree in ITSs. As Regian et al. (1996) have emphasized, our ability to represent human cognition has gained considerable potency with advances made in cognitive science over the last ten years, thereby enhancing substantially the promise and capabilities of ITSs.



Are We Getting Anywhere? Effectiveness.

After more than twenty, or perhaps thirty years (Feurzeig, et al. 1964) of developing ITSs, it may be reasonable to ask how they, or we, are doing. Are they instructionally more effective than what we have? Are they getting better? What data do we have to support the intuitive appeal of ITSs? In a review of 233 empirical evaluations almost entirely made up of standard (non-ITS) computer-based instruction compared with conventional classroom approaches, Kulik (1994) reported an effect size advantage for computer-based approaches of about 0.39 standard deviations, or roughly an improvement of 50th percentile students to about the 65th percentile of achievement. In an attempt to determine the advantages to instruction added by multimedia capabilities, Fletcher (1990) reviewed 47 evaluations of interactive multimedia instruction compared to conventional classroom approaches, and found an effect size advantage for these approaches of about 0.50 standard deviations, or roughly, an improvement of 50th percentile students to about the 69th percentile of achievement. A review performed for this workshop involving 11 ITS evaluations found an effect size advantage for ITSs of about 0.84 standard deviations, or roughly, an improvement of 50th percentile students to about the 79th percentile of achievement. In a recent review covering five evaluations of the SHERLOCK ITS, Gott, Kane and Lesgold (1995), show an overall effect size advantage for SHERLOCK of about 1.05 standard deviations, or roughly, an improvement of 50th percentile students to about the 85th percentile of achievement. We cannot say if we will achieve Bloom's target improvement of two standard deviations, or even exceed it, but the available evidence sugge ts that we are progressing in the right direction.

How Are We Doing?

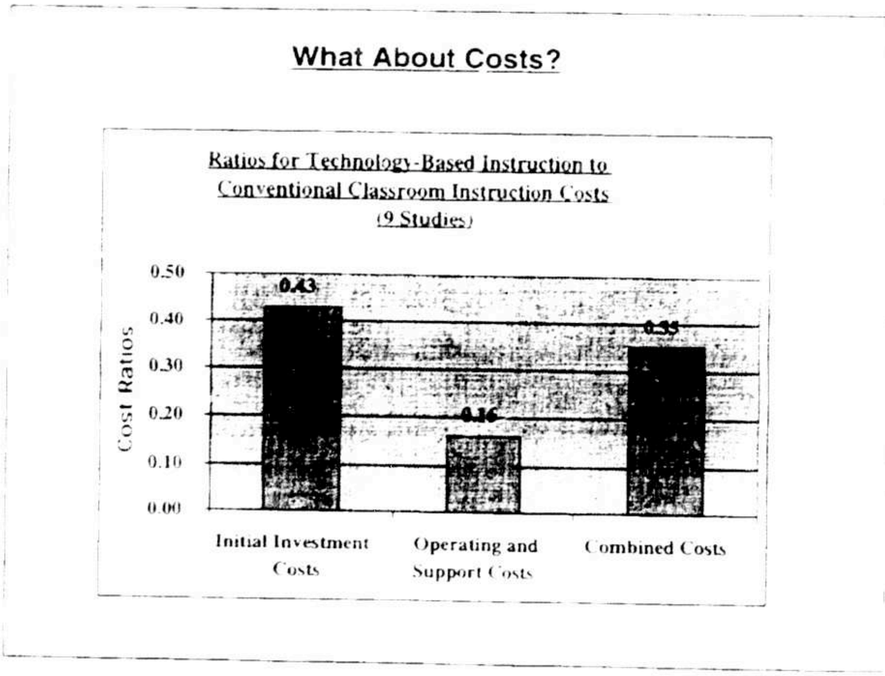


Are We Getting Anywhere? Costs.

Nearly all administrative decision making must consider what is to be gained in the light of what must be given up to get it. Generally, these decisions involve tradeoffs between costs and effectiveness. Decision making in training is no different. Costs must be considered as well as effectiveness.

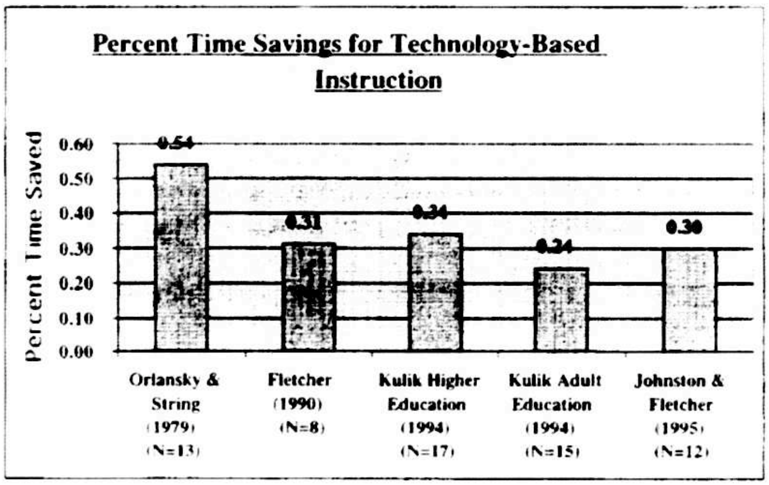
Most ITSs intended for DoD training simulate devices that students must learn to operate or maintain. These devices may cost 1-2 orders of magnitude more than a desktop computer system used to host an ITS. This consideration suggests that favorable cost arguments can be made for ITSs. Unfortunately, little such data exists. However, cost data can be extracted from evaluations that used adequate models of costs to assess the benefits of computer-based instruction using simulated equipment.

A review performed for this presentation found nine evaluations in which the performance of computer-based instruction students trained on simulated equipment and tested on real equipment was found to be at least as good as that of students trained using only real equipment. In fact, the performance of the computer-based instruction students was superior in every case, but the requirement for inclusion in this review was that it only be as good. The ratios of costs for the computer-based instruction to the more conventional instruction using real equipment are shown in three categories; initial investment costs to develop and implement both types of training; operating and support costs to support both types of training once it is in place; and these two cost categories combined. The smaller the ratio then, the better the news for the technology-based training. The results are as shown: 0.43 for initial investment; 0.16 for operating and support; and 0.35 overall.



Are We Getting Anywhere? Time to Learn.

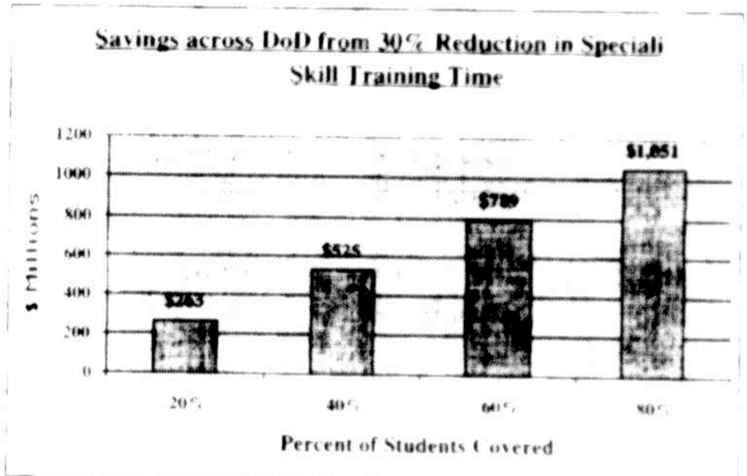
Technology-based instruction should reduce time to reach instructional objectives. By avoiding material the student already knows, concentrating on material yet to be learned, and otherwise keying on each student's pace and style, technology should allow learning to occur more quickly. The reviews of evaluation studies shown in this slide (where the "Ns" refer to numbers of studies reviewed) suggest that these time savings can be obtained and that they average about 30 percent. This is a robust finding in that it appears almost inevitably across many independently performed reviews.



Some Cost Implications of Savings in Time to Learn

To take an example, what would be the cost benefit of reducing time to perform specialized skill training for all DoD students? Specialized skill training is a Congressionally defined category that provides officer and enlisted personnel with skills and knowledge needed to perform specific jobs. It is performed in formally convened military schools. It provides initial training and certification for military occupational specialties such as vehicle mechanics, electronics repair and radar operation. It does not include recruit training, flight training, training performed by operational units or field exercises. In FY 1999 the DoD will provide specialized skill training for about 900,000 students at a cost of about $4.4 billion.

Suppose we were to save 30 percent of the time to provide this training to about 20 percent of the specialized skill students. How much would the DoD save? There is a modestly-careful cost analysis behind this finding, but the end result suggests, as the chart shows, that the DoD would save about $263 million each year. Suppose we were to save 30 percent of the time to provide this training to about 60 pe··ent of the specialized skill students. How much would the DoD save? As the chart shows, the savings would be about $789 million each year. A time savings of 30 percent is not at all unreasonable, as the reviews of time savings show. When a concentrated effort is made to reduce time to train, as for instance, Noja (1987) has done, it is not unreasonable to expect reliable time savings of 50 percent.



The "Thirds"

In summary, all of the assessments of technology-based instruction leave us with "the thirds." Use of technology reduces the cost of instruction by about one-third and, additionally, it either reduces time to reach given instructional objectives by one-third or it increases the achievement of its students by about one-third. The primary payoff for the DoD is, of course, the more rapid preparation of military personnel for operational duties and the increased accessibility of training for sustaining and increasing personal competencies within operational units.

The "Thirds"

Use of technology-based instruction reduces cost of instruction by about 1/3

And

Either ...

- **Reduces time of instruction by about 1/3**

Or ...

- **Increases effectiveness of instruction by about 1/3**

But the REAL Payoff is in Improved Readiness

Some Conclusions

More extensive tailoring of instruction to the needs of individual students obtained through the use of ITSs can only be expected to increase. ITSs may raise the bar for the ultimate effectiveness of technology-based instruction. They may make available far greater efficiencies than we can now obtain from other approaches using technology in instruction. More concentrated attention to the systematic and empirical assessment of ITSs should improve their capabilities and promise. In a formative sense these assessments should improve the design of environments in which individuals learn effectively and efficiently. In a summative sense, they may demonstrate to all stakeholders that the results are worth the effort.

In Conclusion ...

- **ITSs must assert their unique capabilities to raise the ceiling on training effectiveness and efficiency.**
- **Individualization and generative capacity may be key**
  - **We could use more data:**
    - **Formative**
    - **Summative (costs)**

References

1. Atkinson, R. C. and Fletcher, J. D., "Teaching Children to Read with a Computer," *The Reading Teacher*, Vol. 25, 1972, pp. 319-327.
2. Bloom, B. S., "The 2-Sigma Problem: The Search for Methods of Group Instruction as Effective as One-to-One Tutoring" *Educational Researcher*, Vol. 13, No. 6, 1984, pp. 4-16.
3. Brown, J. S., Burton, R. R. and DeKleer, J., "Pedagogical, Natural Language and Knowledge Engineering in SOPHIE I, II and III," in *Intelligent Tutoring Systems*, D. Sleeman and J. S. Brown (eds.), Academic Press, New York, NY, 1982.
4. Carbonell, J. R., "AI in CAI: An Artificial Intelligence Approach to Computer-Assisted Instruction," *IEEE Transactions on Man-Machine Systems*, Vol. 11, 1970, pp. 190-202.
5. Carroll, J. B., "Problems of Measurement Related to the Concept of Learning for Mastery," *Educational Horizons*, Vol. 48, 1970, pp. 71-80.
6. Crowder, N. W., "Automatic Tutoring by Means of Intrinsic Programming," in *Automatic Teaching: The State of the Art*, E. H. Galanter (ed.), John Wiley, New York, NY, 1959.
7. Feurzeig, W., Munter, P., Swets, J. and Breen, M., "Computer-Aided Teaching in Medical Diagnosis," *The Journal of Medical Education*, Vol. 39, 1964, pp. 746-754.
8. Fletcher, J. D., "Artificial Intelligence Applications in Computer-Based Instruction," in *Proc. of ADCIS Summer Meeting*, A. G. Smith (ed.), University of Maine, Portland, ME, 1975.
9. Fletcher, J. D., "Intelligent Training Systems in the Military," in *Defense Applications of Artificial Intelligence: Progress and Prospects*, S. J. Andriole and G. W. Hopple (eds.), Lexington Books, Lexington, MA, 1988.
10. Fletcher, J. D., *The Effectiveness of Interactive Videodisc Instruction in Defense Training and Education*, Institute for Defense Analyses, Alexandria, VA, IDA Paper P-2372 (DTIC No. ADA 228 387), 1990.
11. Fletcher, J. D. and Rockway, M. R., "Computer-Based Training in the Military," in *Military Contributions to Instructional Technology*, J. A. Ellis (ed.) Praeger Publishers, New York, NY, 1986.
12. Gettinger, M., and White, M. A., "Evaluating Curriculum Fit with Class Ability," *Journal of Educational Psychology*, Vol. 72, 1980, pp. 338-344.
13. Gott, S. P., Kane, R. S. and Lesgold, A., *Tutoring for Transfer of Technical Competence*, Armstrong Laboratory, Human Resources Directorate, Brooks AFB, TX (AL/HR-TP-1995-0002), 1995.
14. Kulik, C-L. C., Kulik, J. A. and Shwalb, B., "Effectiveness of Computer-Based Instruction in Adult Education," *Journal of Educational Computing Research*, Vol. 2, 1986, pp. 235-252.
15. Kulik, J. A., "Meta-Analytic Studies of Findings on Computer-Based Instruction," in *Technology Assessment in Education and Training*, E. L. Baker and H. F. O'Neil, Jr. (eds.), Lawrence Erlbaum Associates, Hillsdale, NJ, 1994.
16. Kulik, J .A. and Kulik, C-L. C., "Effectiveness of Computer-Based Education in Colleges," *AEDS Journal*, Vol. 19, 1986, 81-108.
17. Noja, G. P., "DVI and System Integration: A Further Step in ICAI/IMS Technology," in *Advanced Technologies Applied to Training Design*, R. J. Seidel and P.R. Chatelier (eds.), Plenum Press, New York, 1991.

References (Cont'd.)

18. Orlansky, J. and String, J., *Cost-Effectiveness of Computer Based Instruction in Military Training*, Institute for Defense Analyses, Alexandria, Virginia, IDA Paper P-1375, 1979.
19. Regian, W., Seidel, R., Schuler, J. and Radke, P., *Functional Area Analysis of Intelligent Computer-Assisted Instruction*, Report to the Director of Defense Research and Engineering, Training and Personnel Systems Science and Technology Evaluation and Management Committee, Washington, D.C., 1996.
20. Suppes, P., "Modern Learning Theory and the Elementary-School Curriculum," *American Educational Research Journal*, Vol. 1, 1964, pp. 79-93.
21. Suppes, P., Fletcher, J. D., and Zanotti, M., "Performance Models of American Indian Students on Computer-Assisted Instruction in Elementary Mathematics," *Instructional Science*, Vol. 4, 1975, pp. 303-313.
22. Suppes, P., Fletcher, J. D. and Zanotti, M., "Models of Individual Trajectories in Computer-Assisted Instruction for Deaf Students," *Journal of Educational Psychology*, Vol. 68, 1976, pp. 117-127.
23. Suppes, P. and Morningstar, M., *Computer-Assisted Instruction at Stanford 1966-68: Data, Models, and Evaluation of the Arithmetic Programs*, Academic Press, New York, NY, 1972.